

2F04238-PCT

1

DESCRIPTION

COMMUNICATION DEVICE, SIGNAL ENCODING/DECODING METHOD

5 Technical Field

[0001] The present invention relates to a communication
apparatus and signal coding/decoding method for when
speech/audio signals are transmitted in a packet
communication system typified by Internet communication,
10 mobile communication system or the like.

Background Art

[0002] When a speech/audio signal is transmitted using
a packet communication system represented by an Internet
15 communication or mobile communication system, a
compression/coding technology is often used to enhance
transmission efficiency of the speech/audio signal.
Furthermore, with regard to multiplexing of signals, the
smaller the transmission bit rate of each communication
20 terminal, the more communications can be multiplexed,
and therefore for many subscribers to simultaneously
communicate, it is desirable to adopt a technique that
reduces a transmission bit rate of each communication
terminal and enhance the efficiency of channels.

25 [0003] In this respect, there are conventionally
disclosed technologies for reducing a transmission bit
rate in a communication terminal and base station by

acquiring information such as the number of simultaneously accessing users, call loss rate, access waiting time, BER (Bit Error Rate), SIR (Signal Interference Ratio), selecting an appropriate mode from
5 among a plurality of predetermined communication modes according to the information acquired and carrying out communication (e.g., Patent Document 1).

[0004] Furthermore, a technique of detecting the presence/absence of speech of a speaker and controlling
10 a transmission bit rate according to its detection result, is also developed. For example, Non-patent Document 1 discloses a technology of detecting the presence/absence of speech of a speaker, transmitting data coded at a high bit rate for a period during which the speaker is speaking
15 (voiced period), coded at a low bit rate for a period during which the speaker is not speaking (unvoiced period) so as to reduce the overall transmission bit rate (e.g., Non-patent Document 1).

20 Patent Document 1: Japanese Patent Application Laid-Open No. 11-331936

Non-patent Document 1: ANSI/TIA/EIA-96-C, Speech Service Option Standard for Wideband Spread Spectrum Digital Cellular System

25

Disclosure of Invention

Problems to be Solved by the Invention

[0005] However, the above described conventional speech/music coding/decoding method only performs such control as to lower a transmission bit rate when silence continues for a certain time during a conversation as one of elements of the communication environment on the transmitting side and gives no consideration to the operating environment on the receiving side, and therefore it has a problem that efficient transmission is not possible.

10 [0006] It is therefore an object of the present invention to provide a communication apparatus and signal coding/decoding method capable of performing efficient coding on speech/audio signals while maintaining predetermined quality by controlling a transmission bit rate on the transmitting side with the operating environment on the receiving side taken into consideration.

Means for Solving the Problem

20 [0007] The communication apparatus according to the present invention adopts a configuration comprising a transmission mode determining section that determines a transmission mode for controlling a transmission bit rate of a signal transmitted from an apparatus of the communicating party according to a level of ambient noise included in an input signal and transmits the transmission mode to the apparatus of the communicating party and a

25

decoding section that decodes an information source code obtained by coding the input signal at a transmission bit rate corresponding to the transmission mode at the apparatus of the communicating party based on the
5 transmission mode transmitted from the apparatus of the communicating party.

[0008] The communication apparatus of the present invention adopts a configuration comprising a transmission mode determining section that determines
10 a first transmission mode for controlling a transmission bit rate of a signal transmitted from the communication apparatus according to a level of ambient noise included in an input signal of an apparatus of the communicating party and a second transmission mode for controlling a
15 transmission bit rate of an input signal of the communication apparatus based on a level of ambient noise included in the input signal of the communication apparatus and a coding section that performs coding on the input signal at the transmission bit rate
20 corresponding to the second transmission mode and transmits an information source code obtained through the coding and the second transmission mode to the apparatus of the communicating party.

[0009] The communication apparatus according to the
25 present invention adopts a configuration comprising a decoding section that decodes an information source code obtained through coding by an apparatus of the

communicating party, a transmission mode determining section that determines a transmission mode for controlling a transmission bit rate of an input signal according to a level of ambient noise of the signal decoded
5 by the decoding section and a coding section that performs coding on the input signal at a transmission bit rate corresponding to the transmission mode determined by the transmission mode determining section and transmits the information source code obtained through the coding and
10 the transmission mode to the apparatus of the communicating party.

[0010] The communication apparatus according to the present invention adopts a configuration comprising a decoding section that decodes an information source code
15 obtained through coding by an apparatus of the communicating party, a transmission mode determining section that determines a transmission mode for controlling a transmission bit rate of the input signal based on a level of ambient noise included in an input
20 signal and a level of ambient noise of the signal decoded by the decoding section and a coding section that performs coding on the input signal at a transmission bit rate corresponding to the transmission mode determined by the transmission mode determining section and transmits the
25 information source code obtained through the coding and the transmission mode to the apparatus of the communicating party.

[0011] The communication apparatus according to the present invention adopts a configuration comprising a transmission mode determining section that determines a transmission mode for controlling a transmission bit rate of a signal transmitted from an apparatus of the communicating party according to a level of ambient noise included in an input signal and transmits the transmission mode to the apparatus of the communicating party and a decoding section that decodes an information source code obtained by coding the input signal at a transmission bit rate corresponding to the transmission mode by the apparatus of the communicating party based on the transmission mode determined by the transmission mode determining section.

[0012] The signal coding/decoding method according to the present invention is a signal coding/decoding method whereby a first communication apparatus and a second communication apparatus carry out a radio communication, the second communication apparatus transmits an information source code obtained by coding an input signal to the first communication apparatus and the first communication apparatus decodes the information source code, comprising a step by the first communication apparatus of determining a transmission mode for controlling a transmission bit rate of a signal transmitted from the second communication apparatus according to a level of ambient noise included in the

input signal and transmitting the transmission mode to the second communication apparatus, a step by the second communication apparatus of coding the input signal at a transmission bit rate corresponding to the transmission mode determined by the first communication apparatus and transmitting the information source code obtained through the coding to the first communication apparatus and a step by the first communication apparatus of decoding the information source code at the transmission bit rate transmitted from the second communication apparatus.

[0013] The signal coding/decoding method according to the present invention comprises a step of determining a transmission mode for controlling a transmission bit rate of a signal transmitted from an apparatus of the communicating party according to a level of ambient noise included in an input signal and transmitting the transmission mode to the apparatus of the communicating party and a step by the apparatus of the communicating party of decoding an information source code obtained by coding the input signal at a transmission bit rate corresponding to the transmission mode based on the transmission mode transmitted from the apparatus of the communicating party.

[0014] The signal coding/decoding method according to the present invention comprises a step by an apparatus of the communicating party of decoding an information source code obtained through coding, a step of determining

a transmission mode for controlling a transmission bit rate of an input signal according to a level of ambient noise of the decoded signal and a step of coding the input signal at a transmission bit rate corresponding to the
5 determined transmission mode and transmitting the information source code obtained through the coding and the transmission mode to the apparatus of the communicating party.

10 Advantageous Effect of the Invention

[0015] When noise of cars or trains exists on the receiving side, the present invention determines a bit rate on the transmitting side using a masking effect of ambient noise on the receiving side to allow the
15 transmitting side to communicate at a minimum transmission bit rate within a range not influencing human auditory sense, and can thereby substantially improve channel efficiency.

20 Brief Description of Drawings

[0016]

FIG.1 illustrates an auditory masking effect;

FIG.2 is a block diagram showing the configuration of a communication terminal apparatus according to
25 Embodiment 1 of the present invention;

FIG.3 is a block diagram showing the internal configuration of the transmission mode determining

section of the communication terminal apparatus according to the above described embodiment;

FIG.4 is a block diagram showing the internal configuration of the signal coding section of the communication terminal apparatus according to the above described embodiment;

FIG.5 is a block diagram showing the internal configuration of the base layer coding section of the communication terminal apparatus according to the above described embodiment;

FIG.6 is a block diagram showing the internal configuration of the base layer decoding section of the communication terminal apparatus according to the above described embodiment;

FIG.7 is a block diagram showing the internal configuration of the signal decoding section of the communication terminal apparatus according to the above described embodiment;

FIG.8 is a block diagram showing the internal configuration of the signal coding section of the communication terminal apparatus according to the above described embodiment;

FIG.9 is another block diagram showing the internal configuration of the signal decoding section of the communication terminal apparatus according to the above described embodiment;

FIG.10 is a block diagram showing the configuration

of a communication terminal apparatus according to Embodiment 2 of the present invention;

FIG.11 is a block diagram showing the internal configuration of the transmission mode determining
5 section of the communication terminal apparatus according to the above described embodiment;

FIG.12 is a block diagram showing the configuration of a communication apparatus according to Embodiment 3 of the present invention;

10 FIG.13 is a block diagram showing the configuration of a communication terminal apparatus according to Embodiment 4 of the present invention;

FIG.14 is a block diagram showing the internal configuration of the transmission mode determining
15 section of the communication terminal apparatus according to the above described embodiment;

FIG.15 is a block diagram showing the configuration of a communication terminal apparatus according to Embodiment 5 of the present invention;

20 FIG.16 is a block diagram showing the internal configuration of the transmission mode determining section of the communication terminal apparatus according to the above described embodiment;

FIG.17 is a block diagram showing the configuration
25 of a communication terminal apparatus and relay station according to Embodiment 6 of the present invention;

FIG.18 is a block diagram showing the configuration

of the relay station according to the above described embodiment; and

FIG.19 is another block diagram showing the configuration of the relay station according to the above
5 described embodiment.

Best Mode for Carrying Out the Invention

[0017] An audio coding scheme represented by MP3 (Mpeg-1 Audio Layer-3) and AAC (Advanced Audio Coding) realizes
10 efficient coding by using an auditory masking effect and realizing quantization such that quantization errors during coding for each band falls to or below a masking level calculated from an audio signal to be coded. The "auditory masking effect" refers to the phenomenon where
15 the presence of high energy component of a certain frequency "masks" and makes low energy components of neighboring frequencies inaudible.

[0018] FIG.1 illustrates an auditory masking effect. Component B and component C in FIG.1 are masked by component
20 A and component D and cannot be auditorily sensed. Therefore, even when masked components such as component B and component C are reduced a great deal, such a reduction is not perceived. Furthermore, even when a high energy component (large component in the triangular area in
25 FIG.1) is subjected to rough quantization during coding, such a component is characterized in that its errors (quantization errors) are hardly perceptible to the human

ear.

[0019] The present invention applies a relationship between an auditory masking effect which is often used in an audio coding scheme and quantization errors during
5 coding to ambient noise and controls a transmission bit rate based on the masking level of the ambient noise.

[0020] With reference now to the attached drawings, embodiments of the present invention will be explained in detail below.

10 [0021] (Embodiment 1)

Embodiment 1 will explain a speech/music coding/decoding method whereby a transmission mode is determined with an auditory masking effect of ambient noise taken into consideration and a transmission bit
15 rate is controlled in a bidirectional communication between communication terminals.

[0022] FIG.2 is a block diagram showing the configuration of a communication terminal apparatus according to Embodiment 1. In FIG.2, suppose a bidirectional
20 communication is carried out between two communication terminal apparatuses 100 and 150.

[0023] First, the configuration of communication terminal apparatus 100 will be explained. Communication terminal apparatus 100 is mainly constructed of
25 transmission mode determining section 101, signal coding section 102 and signal decoding section 103.

[0024] Transmission mode determining section 101

detects ambient noise included in the background of a speech/audio signal in an input signal and determines a transmission mode for controlling a transmission bit rate of a signal transmitted from communication terminal apparatus 150, which is the communication terminal of the communicating party, according to the level of ambient noise. Transmission mode determining section 101 outputs information indicating the determined transmission mode (hereinafter referred to as "transmission mode information") to transmission path 110 and signal decoding section 103. In an example of this embodiment, suppose that one transmission bit rate is selected from two or more predetermined transmission bit rates and the transmission mode information can take three types of transmission bit rate values; bitrate 1, bitrate 2, bitrate 3 (bitrate 3 < bitrate 2 < bitrate 1).

[0025] Signal coding section 102 performs coding on the input signal which is a speech/audio signal according to the transmission mode information transmitted from communication terminal apparatus 150 through transmission path 110 and outputs the obtained coded information to transmission path 110.

[0026] Signal decoding section 103 decodes coded information transmitted from communication terminal apparatus 150 through transmission path 110 and outputs the obtained signal as an output signal. Signal decoding

section 103 compares the transmission mode information included in the coded information output from transmission path 110 with the transmission mode information obtained from transmission mode determining section 101 with a transmission delay taken into consideration, and can thereby detect transmission errors. To be more specific, when the transmission mode information obtained from transmission mode determining section 101 with a transmission delay taken into consideration is different from the transmission mode information included in the coded information output from transmission path 110, signal decoding section 103 decides that a transmission error has occurred in transmission path 110. Furthermore, it is also possible to adopt a technique whereby signal coding section 152 of communication terminal apparatus 150 does not integrate the transmission mode information with the coded information, while signal decoding section 103 decodes the coded information output from transmission path 110 using the transmission mode information obtained from transmission mode determining section 101.

[0027] Next, the configuration of communication terminal apparatus 150 will be explained. Communication terminal apparatus 150 is mainly constructed of transmission mode determining section 151, signal coding section 152 and signal decoding section 153.

[0028] Transmission mode determining section 151 is fed

an input signal, detects ambient noise included in the background of a speech/audio signal and determines a transmission mode for controlling a transmission bit rate of a signal transmitted from communication terminal apparatus 100 according to the level of ambient noise. 5 Next, transmission mode determining section 151 outputs the transmission mode information indicating the determined transmission mode to transmission path 110 and signal decoding section 153.

10 [0029] Signal coding section 152 is fed the transmission mode information transmitted from communication terminal apparatus 100 through transmission path 110, performs coding on the input signal which is a speech/audio signal according to the transmission mode information and 15 outputs the obtained coded information to transmission path 110.

[0030] Signal decoding section 153 is fed the coded information transmitted from communication terminal apparatus 100 through transmission path 110 and the 20 transmission mode information obtained from transmission mode determining section 151, decodes the coded information and outputs the obtained signal as an output signal. By comparing the transmission mode information included in the coded information output from 25 transmission path 110 with the transmission mode information obtained from the transmission mode determining section 151 with a transmission delay taken

into consideration, signal decoding section 153 can detect transmission errors. To be more specific, when the transmission mode information obtained from transmission mode determining section 151 with a transmission delay taken into consideration is different from the transmission mode information included in the coded information output from transmission path 110, signal decoding section 153 decides that a transmission error has occurred in transmission path 110. Furthermore, it is also possible to adopt a technique whereby signal coding section 102 of communication terminal apparatus 100 does not integrate the transmission mode information with the coded information and signal decoding section 153 decodes the coded information output from transmission path 110 using the transmission mode information obtained from transmission mode determining section 151.

[0031] Next, the internal configuration of transmission mode determining section 101 in FIG.2 will be explained using FIG.3. The configuration of transmission mode determining section 151 in FIG.2 is the same as that of transmission mode determining section 101.

[0032] Transmission mode determining section 101 is mainly constructed of masking level calculation section 301 and transmission mode decision section 302.

[0033] Masking level calculation section 301 calculates a masking level from the input signal and outputs the

calculated masking level to transmission mode decision section 302.

[0034] Transmission mode decision section 302 compares the masking level output from masking level calculation section 301 with a predetermined threshold and determines a transmission bit rate based on the comparison result. To be more specific, when the level of ambient noise existing in communication terminal apparatus 100 detected by communication terminal apparatus 100 is large and its masking level is large, the transmission bit rate is decreased. This is based on a principle that a quantization error of the coded information transmitted from communication terminal apparatus 150 is masked to a certain extent through an auditory masking effect of ambient noise, and, therefore, even when transmission bit rate is lowered at communication terminal apparatus 150, a decoded signal is obtained in equal auditory quality to the case where the transmission bit rate is not lowered. On the other hand, when the level of ambient noise existing on the communication terminal apparatus 100 side detected by communication terminal apparatus 100 is small, the quantization error of the coded information transmitted from communication terminal apparatus 150 is not masked by the auditory masking effect of ambient noise, and therefore the transmission bit rate is increased.

[0035] Transmission mode decision section 302 outputs the transmission mode information indicating the

determined transmission mode to transmission path 110 and signal decoding section 103.

[0036] Here, the processing of masking level calculation section 301 and transmission mode decision section 302 in the case will be explained where a method is adopted whereby transmission mode determining section 101 calculates a maximum value and minimum value of the power value of the input signal for a predetermined period of time (e.g., a certain period of approximately 5 seconds to 10 seconds), decides the level of ambient noise included in the input signal from the maximum value and minimum value and the bit rate is controlled according to the level. Here, a case where processing of deciding and outputting the level of ambient noise is carried out every time a frame is processed will be explained, but, in addition to this, it is also possible to perform subsequent processing with pressing of a button by the user of the communication terminal as a trigger or perform subsequent processing at certain time intervals. Furthermore, it is also possible to detect the level of ambient noise at certain time intervals and perform subsequent processing when the difference between the detected level of ambient noise and the previous detected level exceeds a predetermined threshold.

[0037] First, the processing of masking level calculation section 301 will be explained. Masking level calculation section 301 divides the input signal into

groups of N samples (N : natural number), regards each interval as 1 frame and performs processing in frame units. Hereinafter, the input signal to be coded will be expressed as x_n ($n=0, \dots, N-1$).

5 [0038] Furthermore, masking level calculation section 301 includes buffers buf_i ($i=0, \dots, N_i-1$). Here, N_i denotes a predetermined non-negative integer, which depends on the number of samples N of 1 frame and when
10 a 1-frame interval is on the order of approximately 20 milliseconds, it is confirmed that desired performance can be obtained when N_i is a value on the order of 100 to 500.

[0039] Next, masking level calculation section 301 will calculate frame power P_{frame} of the frame to be processed
15 from Equation 1 below:

[0040]

[Equation 1]

$$P_{\text{frame}} = \sum_{n=0}^{N-1} |x_n|^2 \quad \dots (1)$$

Next, masking level calculation section 301
20 substitutes frame power P_{frame} calculated from Equation 1 into buffer buf_{N_i-1} .

[0041] Next, masking level calculation section 301 calculates minimum value $P_{\text{frame_MIN}}$ and maximum value $P_{\text{frame_MAX}}$ of frame power P_{frame} in an i interval (interval
25 length N_i) and outputs $P_{\text{frame_MIN}}$, $P_{\text{frame_MAX}}$ to transmission mode decision section 302.

[0042] Next, masking level calculation section 301 updates buffer buf_i according to Equation 2 below.

[0043]

[Equation 2]

$$5 \quad buf_i = buf_{i+1} \quad (i = 0, \dots, N_i - 2) \quad \dots (2)$$

This is the explanation of the processing by masking level calculation section 301 in FIG.3.

[0044] Next, the processing of transmission mode decision section 302 will be explained. Transmission mode decision section 302 determines transmission mode information mode from $Pframe_{MIN}$, $Pframe_{MAX}$ output from masking level calculation section 301, according to Equation 3 below:

[0045]

15 [Equation 3]

$$Mode = \begin{cases} bitrate_1 & (Th_0 \leq Pframe_{MAX} / Pframe_{MIN}) \\ bitrate_2 & (Th_1 \leq Pframe_{MAX} / Pframe_{MIN} < Th_0) \dots (3) \\ bitrate_3 & (Pframe_{MAX} / Pframe_{MIN} < Th_1) \end{cases}$$

where Th_0 and Th_1 ($Th_0 < Th_1$) are constants predetermined by a preliminary experiment based on a auditory masking effect of ambient noise.

20 [0046] Hereinafter, the preliminary experiment for calculating Th_0 and Th_1 will be briefly explained. Here, a coding method used when mode is bitrate 1 is referred to as coding method A, and a signal obtained by decoding information coded by coding method A is referred to as
25 decoded signal A. Likewise, a coding method used when

mode is bitrate 2 is referred to as coding method B, and a signal obtained by decoding information coded by coding method B is referred to as decoded signal B. Furthermore, a coding method used when mode is bitrate 3 is referred to as coding method C and a signal obtained by decoding information coded by coding method C is referred to as decoded signal C.

[0047] When average noise (e.g., white noise) is gradually added to decoded signal A and decoded signal B such that its level is gradually increased, suppose the noise level when noise-added decoded signal A becomes auditorily equal to noise-added decoded signal B is Th_0 . Likewise, suppose noise level when noise-added decoded signal A becomes auditorily equal to noise-added decoded signal C is Th_1 . In this way, Th_0 and Th_1 are experimentally determined using the masking effect of noise.

[0048] Next, transmission mode decision section 302 outputs the transmission mode information to transmission path 110 and signal decoding section 103.

[0049] This is the explanation of the internal configuration of transmission mode determining section 101 in FIG.2.

[0050] Next, the configuration of signal coding section 102 in FIG.2 will be explained using FIG.4. Note that the configuration of signal coding section 152 in FIG.2 is the same as that of signal coding section 102.

[0051] Here, a case will be described with this embodiment where a speech/audio signal is coded/decoded using a three-layer speech coding/decoding method made up of one base layer and two enhancement layers. However, 5 the present invention places no restrictions on the number of layers and the present invention is also applicable to cases where a speech/audio signal is coded/decoded using a layered speech coding/decoding method having four or more layers.

10 [0052] The "layered speech coding method" is a method in which a plurality of speech coding methods whereby a residual signal (difference between an input signal in a lower layer and a decoded signal in a lower layer) is coded and the coded information is output exist in 15 a higher layer, forming a layered structure. Furthermore, the "layered speech decoding method" is a method in which a plurality of speech decoding methods whereby a residual signal is decoded exist in a higher layer, forming a layered structure. Here, suppose the speech coding/decoding 20 method which exists in the lowest layer is a base layer. Furthermore, suppose a speech coding/decoding method which exists in a higher layer than the base layer is an enhancement layer. Hereinafter, the coding section and the decoding section in the base layer are referred 25 to as a base layer coding section and a base layer decoding section respectively and the coding section and the decoding section in an enhancement layer are referred

to as an enhancement layer coding section and an enhancement layer decoding section respectively.

[0053] Signal coding section 102 is mainly constructed of transmission bit rate control section 401, control
5 switches 402 to 405, base layer coding section 406, base layer decoding section 407, addition sections 408 and 411, first enhancement layer coding section 409, first enhancement layer decoding section 410, second enhancement layer coding section 412 and coded
10 information integration section 413.

[0054] An input signal is input to base layer coding section 406 and control switch 402. Furthermore, transmission mode information is input to transmission bit rate control section 401.

15 [0055] Transmission bit rate control section 401 performs ON/OFF control of control switches 402 to 405 according to the input transmission mode information. To be more specific, when the transmission mode information is bitrate 1, transmission bit rate control
20 section 401 sets all control switches 402 to 405 to ON. Furthermore, when the transmission mode information is bitrate 2, transmission bit rate control section 401 sets control switches 402 and 403 to ON and sets control switches 404 and 405 to OFF. Furthermore, when the transmission
25 mode information is bitrate 3, transmission bit rate control section 401 sets all control switches 402 to 405 to OFF. In this way, transmission bit rate control

section 401 performs ON/OFF control of the control switches according to the transmission mode information and a combination of coding sections used for coding of an input signal is thereby determined. Note that the transmission mode information is output from transmission bit rate control section 401 to coded information integration section 413.

[0056] Base layer coding section 406 performs coding on the input signal and outputs an information source code obtained through the coding (hereinafter referred to as "base layer information source code") to control switch 403 and coded information integration section 413. The internal configuration of base layer coding section 406 will be described later.

[0057] When control switch 403 is ON, base layer decoding section 407 decodes the base layer information source code output from base layer coding section 406 and outputs the obtained decoded signal (hereinafter referred to as "base layer decoded signal") to addition section 408. When control switch 403 is OFF, base layer decoding section 407 performs no operation. The internal configuration of base layer decoding section 407 will be described later.

[0058] When control switches 402 and 403 are ON, addition section 408 adds a signal obtained by inverting the polarity of the base layer decoded signal output from base layer decoding section 407 to the input signal and outputs a first residual signal, which is the addition

result, to first enhancement layer coding section 409 and control switch 404. When control switches 402 and 403 are OFF, addition section 408 performs no operation.

[0059] When control switches 402 and 403 are ON, first
5 enhancement layer coding section 409 performs coding on the first residual signal output from addition section 408 and outputs the information source code obtained through the coding (hereinafter referred to as "first enhancement layer information source code") to control
10 switch 405 and coded information integration section 413. When control switches 402 and 403 are OFF, first enhancement layer coding section 409 performs no operation.

[0060] When control switch 405 is ON, first enhancement
15 layer decoding section 410 decodes the first enhancement layer information source code output from first enhancement layer coding section 409 and outputs the obtained decoded signal through the decoding (hereinafter referred to as "first enhancement layer decoded signal")
20 to addition section 411. When control switch 405 is OFF, first enhancement layer decoding section 410 performs no operation.

[0061] When control switches 404 and 405 are ON, addition
25 section 411 adds a signal obtained by inverting the polarity of the output signal of first enhancement layer decoding section 410 to the first residual signal and outputs a second residual signal, which is the addition

result, to second enhancement layer coding section 412. When control switches 404 and 405 are OFF, addition section 411 performs no operation.

[0062] When control switches 404 and 405 are ON, second
5 enhancement layer coding section 412 performs coding on the second residual signal output from addition section 411 and outputs the information source code obtained through the coding (hereinafter referred to as "second enhancement layer information source code") to coded
10 information integration section 413. When control switches 404 and 405 are OFF, second enhancement layer coding section 412 performs no operation.

[0063] Coded information integration section 413 integrates the transmission mode information output from
15 transmission bit rate control section 401, base layer information source code output from base layer coding section 406, first enhancement layer information source code output from first enhancement layer coding section 409 and second enhancement layer information source code
20 output from second enhancement layer coding section 412, and outputs the integrated coded information to transmission path 110.

[0064] This is the explanation of the configuration of signal coding section 102 using FIG.4. So far, signal
25 coding section 102 has been explained under the condition that the transmission mode information is always input to transmission bit rate control section 401 during

processing of each frame, but, when the transmission mode information is not input to transmission bit rate control section 401, it is also possible to use transmission mode information of previous input by, for example, storing
5 the previously input transmission mode information in the buffer in transmission bit rate control section 401.

[0065] Next, the configuration of base layer coding section 406 in FIG.4 will be explained using FIG.5. This embodiment will explain a case where base layer coding
10 section 406 performs CELP type speech coding.

[0066] Pre-processing section 501 performs high pass filter processing for removing a DC component, wave shaping processing which will lead to performance improvement of subsequent coding processing and
15 pre-emphasis processing on a signal of an input sampling frequency and outputs a signal (Xin) after these processing to LPC analysis section 502 and addition section 505.

[0067] LPC analysis section 502 performs a linear
20 predictive analysis using Xin and outputs the analysis result (linear predictive coefficient) to LPC quantization section 503. LPC quantization section 503 performs quantization processing on the linear predictive coefficient (LPC) output from LPC analysis section 502
25 and outputs the quantization LPC to synthesis filter 504 and outputs a code (L) indicating the quantization LPC to multiplexing section 514.

[0068] Synthesis filter 504 performs filter synthesis on an excitation vector output from addition section 511 which will be described later using a filter coefficient based on the quantization LPC, thereby generating a
5 composite signal and outputting the composite signal to addition section 505.

[0069] Addition section 505 adds a signal obtained by inverting the polarity of the composite signal to X_{in} , thereby calculating an error signal and outputting the
10 error signal to auditory weighting section 512.

[0070] Adaptive excitation codebook 506 stores excitation vectors output in the past from addition section 511 in a buffer, extracts samples corresponding to 1 frame from a past excitation vector identified by
15 a signal output from parameter determining section 513 as an adaptive excitation vector and outputs it to multiplication section 509.

[0071] Quantization gain generation section 507 outputs a quantization adaptive excitation gain and quantization
20 fixed excitation gain identified by the signal output from parameter determining section 513 to multiplication section 509 and multiplication section 510 respectively.

[0072] Fixed excitation codebook 508 outputs a fixed excitation vector obtained by multiplying a pulse
25 excitation vector having a shape identified by the signal output from parameter determining section 513 by a spreading vector to multiplication section 510.

[0073] Multiplication section 509 multiplies the adaptive excitation vector output from adaptive excitation codebook 506 by the quantization adaptive excitation gain output from quantization gain generation section 507 and outputs the multiplication result to addition section 511. Multiplication section 510 multiplies the fixed excitation vector output from fixed excitation codebook 508 by the quantization fixed excitation gain output from quantization gain generation section 507 and outputs the multiplication result to addition section 511.

[0074] Addition section 511 is fed the gain-multiplied adaptive excitation vector and fixed excitation vector from multiplication section 509 and multiplication section 510 respectively, adds up these vectors and outputs an excitation vector which is the addition result to synthesis filter 504 and adaptive excitation codebook 506. The excitation vector input to adaptive excitation codebook 506 is stored in a buffer.

[0075] Auditory weighting section 512 performs auditory weighting on the error signal output from addition section 505 and outputs the auditory weighting result as coding distortion to parameter determining section 513.

[0076] Parameter determining section 513 selects an adaptive excitation vector, fixed excitation vector and quantization gain that minimize coding distortion output from auditory weighting section 512 from adaptive

excitation codebook 506, fixed excitation codebook 508 and quantization gain generation section 507 respectively and outputs adaptive excitation vector code (A), fixed excitation vector code (F) and excitation gain code (G) indicating the selection result to multiplexing section 514.

[0077] Multiplexing section 514 is fed code (L) indicating the quantization LPC from LPC quantization section 503, is fed code (A) indicating the adaptive excitation vector, code (F) indicating the fixed excitation vector and code (G) indicating the excitation gain from parameter determining section 513 and multiplexes these information and outputs the multiplexing result as a base layer information source code.

[0078] This is the explanation of the internal configuration of base layer coding section 406 in FIG.4.

[0079] The internal configurations of first enhancement layer coding section 409 and second enhancement layer coding section 412 in FIG.4 are the same as that of base layer coding section 406 and are different in only the type of signal input and the type of information source code output, and therefore explanations thereof will be omitted.

[0080] Next, the internal configuration of base layer decoding section 407 in FIG.4 will be explained using FIG.6. Here, a case where base layer decoding section

407 carries out CELP type speech decoding will be explained.

[0081] In FIG.6, a base layer information source code input to base layer decoding section 407 is separated
5 by demultiplexing section 601 into individual codes (L, A, G, F). The separated LPC code (L) is output to LPC decoding section 602, the separated adaptive excitation vector code (A) is output to adaptive excitation codebook 605, the separated excitation gain code (G) is output
10 to quantization gain generation section 606 and the separated fixed excitation vector code (F) is output to fixed excitation codebook 607.

[0082] LPC decoding section 602 decodes quantization LPC from the code (L) output from demultiplexing section 601
15 and outputs it to synthesis filter 603.

[0083] Adaptive excitation codebook 605 extracts samples corresponding to 1 frame from a past excitation vector specified by the code (A) output from demultiplexing section 601 as an adaptive excitation
20 vector and outputs it to multiplication section 608.

[0084] Quantization gain generation section 606 decodes the quantization adaptive excitation gain and quantization fixed excitation gain specified by the excitation gain code (G) output from demultiplexing
25 section 601 and outputs the decoding results to multiplication section 608 and multiplication section 609.

[0085] Fixed excitation codebook 607 generates a fixed excitation vector specified by the code (F) output from demultiplexing section 601 and outputs the fixed excitation vector to multiplication section 609.

5 [0086] Multiplication section 608 multiplies the adaptive excitation vector by the quantization adaptive excitation gain and outputs the multiplication result to addition section 610. Multiplication section 609 multiplies the fixed excitation vector by the
10 quantization fixed excitation gain and outputs the multiplication result to addition section 610.

[0087] Addition section 610 adds up the gain-multiplied adaptive excitation vector and fixed excitation vector output from multiplication sections 608, 609, generates
15 an excitation vector and outputs it to synthesis filter 603 and adaptive excitation codebook 605.

[0088] Synthesis filter 603 performs filter synthesis of the excitation vector output from addition section 610 using the filter coefficient decoded by LPC decoding
20 section 602 and outputs a composite signal to post-processing section 604.

[0089] Post-processing section 604 performs processing of improving subjective quality of speech such as formant emphasis and pitch emphasis or processing of improving
25 subjective quality of stationary noise on the signal output from synthesis filter 603 and outputs the processed signal as base layer decoded information.

[0090] This is the explanation of the internal configuration of base layer decoding section 407 in FIG. 4.

[0091] The internal configuration of first enhancement layer decoding section 410 in FIG. 4 is the same as the
5 internal configuration of base layer decoding section 407 and is different only in the type of information source code input and the type of signal output, and therefore explanations thereof will be omitted.

[0092] Next, the configuration of signal decoding
10 section 103 in FIG. 2 will be explained using FIG. 7. The configuration of signal decoding section 153 in FIG. 2 is the same as the configuration of signal decoding section 103.

[0093] Signal decoding section 103 is mainly constructed
15 of transmission bit rate control section 701, base layer decoding section 702, first enhancement layer decoding section 703, second enhancement layer decoding section 704, control switches 705 and 706 and addition sections 707 and 708.

[0094] Transmission bit rate control section 701
20 controls ON/OFF of control switches 705 and 706 according to transmission mode information included in received coded information. To be more specific, when the transmission mode information is bitrate 1, transmission
25 bit rate control section 701 sets both control switches 705 and 706 to ON. Furthermore, when the transmission mode information is bitrate 2, transmission bit rate

control section 701 sets control switch 705 to ON and sets control switch 706 to OFF. Furthermore, when the transmission mode information is bitrate 3, transmission bit rate control section 701 sets both control switches 705 and 706 to OFF. Furthermore, transmission bit rate control section 701 separates the received coded information into the base layer information source code, first enhancement layer information source code and second enhancement layer information source code included therein, outputs the base layer information source code to base layer decoding section 702, outputs the first enhancement layer information source code to control switch 705 and outputs the second enhancement layer information source code to control switch 706.

[0095] Base layer decoding section 702 decodes the base layer information source code output from transmission bit rate control section 701, generates a base layer decoded signal and outputs it to addition section 708.

[0096] When control switch 705 is ON, first enhancement layer decoding section 703 decodes the first enhancement layer information source code output from transmission bit rate control section 701, generates a first enhancement layer decoded signal and outputs it to addition section 707. When control switch 705 is OFF, first enhancement layer decoding section 703 performs no operation.

[0097] When control switch 706 is ON, second enhancement

layer decoding section 704 decodes the second enhancement layer information source code output from transmission bit rate control section 701, generates a second enhancement layer decoded signal and outputs it to addition section 707. When control switch 706 is OFF, second enhancement layer decoding section 704 performs no operation.

[0098] When control switches 705 and 706 are ON, addition section 707 adds up the second enhancement layer decoded signal output from second enhancement layer decoding section 704 and the first enhancement layer decoded signal output from first enhancement layer decoding section 703, and outputs the signal after the addition to addition section 708. Furthermore, when control switch 706 is OFF and control switch 705 is ON, addition section 707 outputs the first enhancement layer decoded signal output from first enhancement layer decoding section 703 to addition section 708. When control switches 705 and 706 are OFF, addition section 707 performs no operation.

[0099] Addition section 708 adds up the base layer decoded signal output from base layer decoding section 702 and the output signal of addition section 707 and outputs the signal after the addition as an output signal. Furthermore, when control switches 705 and 706 are OFF, addition section 708 outputs the base layer decoded signal output from base layer decoding section 702 as an output signal.

[0100] This is the explanation of the configuration of signal decoding section 103 in FIG.2.

[0101] Note that the internal configurations of base layer decoding section 702, first enhancement layer
5 decoding section 703 and second enhancement layer decoding section 704 in FIG.7 are the same as the internal configuration of base layer decoding section 407 in FIG.4 and are only different in the type of signal input and the type of information source code output, and therefore
10 explanations thereof will be omitted.

[0102] Here, as the coding/decoding method for signal coding section 102 and signal decoding section 103, it is also possible to apply a configuration whereby coding/decoding is performed by switching between a
15 plurality of coding/decoding methods of different bit rates. Hereinafter, the configurations of signal coding section 102 and signal decoding section 103 in this case will be explained using FIG.8 and FIG.9.

[0103] This embodiment will explain the case where
20 speech/audio signals are coded/decoded using three types of speech coding/decoding methods. However, the present invention places no limit on the number of coding/decoding methods and the present invention is also applicable to cases where speech/audio signals are coded/decoded using
25 speech coding/decoding methods of four or more different types of bit rates.

[0104] FIG.8 is a block diagram showing the internal

configuration of signal coding section 102. Signal coding section 102 is mainly constructed of transmission bit rate control section 801, control switches 802 and 803, signal coding sections 804 to 806 and coded
5 information integration section 807.

[0105] An input signal is input to control switch 802. Furthermore, transmission mode information is input to transmission bit rate control section 801.

[0106] Transmission bit rate control section 801
10 controls switching of control switches 802 and 803 according to the input transmission mode information. To be more specific, when the transmission mode information is bitrate 1, transmission bit rate control section 801 connects both control switches 802 and 803
15 to signal coding section 804. Furthermore, when the transmission mode information is bitrate 2, transmission bit rate control section 801 connects both control switches 802 and 803 to signal coding section 805. Furthermore, when the transmission mode information is
20 bitrate 3, transmission bit rate control section 801 connects both control switches 802 and 803 to signal coding section 806. Thus, transmission bit rate control section 801 controls switching of the control switches according to the transmission mode information to thereby determine
25 a coding section to be used for coding of the input signal. The transmission mode information is output from transmission bit rate control section 801 to coded

information integration section 807.

[0107] Signal coding section 804 performs coding on the input signal using a coding method corresponding to bitrate 1 and outputs the information source code obtained
5 through coding to coded information integration section 807 through control switch 803.

[0108] Signal coding section 805 performs coding on the input signal using a coding method corresponding to bitrate 2 and outputs the information source code obtained
10 through coding to coded information integration section 807 through control switch 803.

[0109] Signal coding section 806 performs coding on the input signal using a coding method corresponding to bitrate 3 and outputs the information source code obtained
15 through coding to coded information integration section 807 through control switch 803.

[0110] Coded information integration section 807 integrates the transmission mode information output from transmission bit rate information control section 801
20 and the information source code output from switch 803 and outputs the integrated coded information to transmission path 110.

[0111] This is the explanation of the configuration of signal coding section 102 using FIG.8. The above
25 described case has been explained under the condition that transmission mode information is always input to transmission bit rate control section 801 every time a

frame is processed, but, when the transmission mode information is not input to transmission bit rate control section 801, it is also possible to use previously input transmission mode information by, for example, storing
5 the previously input transmission mode information in a buffer of transmission bit rate control section 801.

[0112] The internal configurations of signal coding sections 804 to 806 in FIG.8 are the same as that of base layer coding section 406 in FIG.4 and are only different
10 in the type of signals input and the type of information source code output, and therefore explanations thereof will be omitted.

[0113] FIG.9 is a block diagram showing the internal configuration of signal decoding section 103. Signal
15 decoding section 103 is mainly constructed of transmission bit rate control section 901, control switches 902 and 903 and signal decoding sections 904 to 906.

[0114] Coded information is input to transmission bit
20 rate control section 901.

[0115] Transmission bit rate control section 901 controls switching of control switches 902 and 903 according to transmission mode information included in received coded information. To be more specific, when
25 the transmission mode information is bitrate 1, transmission bit rate control section 901 connects both control switches 902 and 903 to signal decoding section

904. Furthermore, when the transmission mode information is bitrate 2, transmission bit rate control section 901 connects both control switches 902 and 903 to signal decoding section 905. Furthermore, when the
- 5 transmission mode information is bitrate 3, transmission bit rate control section 901 connects both control switches 902 and 903 to signal decoding section 906. Transmission bit rate control section 901 also outputs a received information source code to control switch 902.
- 10 [0116] Signal decoding section 904 decodes the information source code input through control switch 902 using a decoding method corresponding to bitrate 1 and outputs the output signal obtained through the decoding through control switch 903.
- 15 [0117] Signal decoding section 905 decodes the information source code input through control switch 902 using a decoding method corresponding to bitrate 2 and outputs the output signal obtained through the decoding through control switch 903.
- 20 [0118] Signal decoding section 906 decodes the information source code input through control switch 902 using a decoding method corresponding to bitrate 3 and outputs the output signal obtained through the decoding through control switch 903.
- 25 [0119] This is the explanation of the configuration of signal decoding section 103 using FIG.9.
- [0120] The internal configurations of signal decoding

sections 904 to 906 in FIG.9 are the same as the internal configuration of base layer decoding section 407 in FIG.4 and are only different in the type of information source code input and the type of signal output and explanations thereof will be omitted.

[0121] Thus, it is possible to perform efficient coding of speech/audio signals by controlling a transmission bit rate on the transmitting side according to the masking level of ambient noise with the masking effect of ambient noise on the receiving side taken into consideration.

[0122] (Embodiment 2)

Here, the above described speech coding method such as CELP uses a speech excitation/vocal tract model, and can thereby perform efficient coding about human speech, but cannot perform efficient coding about components other than human speech such as ambient noise existing in the background. Therefore, when ambient noise exists on the transmitting side, in order to perform coding on speech/audio signals including ambient noise on the transmitting side with equal quality to the case where no ambient noise exists, more bits are required than when no ambient noise exists on the transmitting side.

[0123] Embodiment 2 will explain a case where a transmission bit rate is controlled with not only ambient noise on the receiving side but also ambient noise on the transmitting side taken into consideration.

[0124] FIG.10 is a block diagram showing the

configuration of a communication terminal apparatus according to Embodiment 2 of the present invention. In communication terminal apparatuses 1000 and 1050 shown in FIG.10, components common to those of communication terminal apparatuses 100 and 150 shown in FIG.2 are assigned the same reference numerals as those in FIG.2 and explanations thereof will be omitted.

[0125] When communication terminal apparatus 1000 in FIG.10 is compared to communication terminal apparatus 100 in FIG.2, the operation of transmission mode determining section 1001 differs from that of transmissionmode determining section 101. Furthermore, when communication terminal apparatus 1050 in FIG.10 is compared to communication terminal apparatus 150 in FIG.2, the operation of transmission mode determining section 1051 differs from that of transmission mode determining section 151.

[0126] Transmission mode determining section 1001 detects ambient noise included in the background of a speech/audio signal in an input signal, determines a transmission mode for controlling a transmission bit rate of a signal transmitted from communication terminal apparatus 1050, which is a communication terminal of a communicating party, according to the level of ambient noise and outputs transmission mode information indicating the determined transmission mode to transmission path 110. Furthermore, transmission mode

determining section 1001 determines a transmission mode for controlling a transmission bit rate when performing coding/decoding based on the level of ambient noise in an input signal and transmission mode information transmitted from communication terminal apparatus 1050 through transmission path 110 and outputs transmission mode information indicating the determined transmission mode to signal coding section 102 and signal decoding section 103.

[0127] Next, the internal configuration of transmission mode determining section 1001 in FIG.10 will be explained using FIG.11. Transmission mode determining section 1001 is mainly constructed of masking level calculation section 1101 and transmission mode decision section 1102. Here, a case where processing of deciding and outputting the level of ambient noise every time each frame is processed is performed will be explained. In addition to this, it is also possible to carry out subsequent processing with pressing of a button by the user of a communication terminal or the like as a trigger or carry out subsequent processing at predetermined time intervals.

[0128] As in the case of masking level calculation section 301 in FIG.3, masking level calculation section 1101 calculates a masking level from an input signal and outputs the calculated masking level to transmission mode decision section 1102.

[0129] Transmission mode decision section 1102 determines a transmission mode for controlling a transmission bit rate with ambient noise on the transmitting side taken into consideration based on the result of a comparison between the masking level output from masking level calculation section 1101 and a predetermined threshold and outputs information indicating the determined transmission mode (hereinafter referred to as "first transmission mode information") to transmission path 110. Furthermore, transmission mode decision section 1102 determines a transmission mode for controlling a transmission bit rate with ambient noise on the transmitting side and the receiving side taken into consideration based on the first transmission mode information and transmission mode information transmitted from communication terminal apparatus 1050 through transmission path 110 (hereinafter referred to as "second transmission mode information") and outputs information indicating the determined transmission mode (hereinafter referred to as "third transmission mode information") to signal coding section 102 and signal decoding section 103.

[0130] Here, the processing of transmission mode decision section 1102 in the case of adopting a method whereby transmission mode determining section 1001 calculates a maximum value and a minimum value of the power value of an input signal for a predetermined period,

decides the level of ambient noise included in an input signal from the maximum value and minimum value and controls the bit rate according to the level will be explained.

5 [0131] First, transmission mode decision section 1102 determines first transmission mode information $Mode'_1$ from $Pframe_{MIN}$, $Pframe_{MAX}$ output from masking level calculation section 1101 according to Equation 4 below:

[0132]

10 [Equation 4]

$$Mode'_1 = \begin{cases} bitrate_{high} & (Th'_0 \leq Pframe'_{MAX} / Pframe'_{MIN}) \\ bitrate_{low} & (Pframe'_{MAX} / Pframe'_{MIN} < Th'_0) \end{cases} \dots (4)$$

where Th'_0 is a constant predetermined based on an auditory masking effect of ambient noise through an experiment similar to the preliminary experiment
15 explained in Embodiment 1.

[0133] Next, transmission mode decision section 1102 outputs first transmission mode information $Mode'_1$ to transmission path 110.

[0134] Furthermore, transmission mode decision section
20 1102 calculates third transmission mode information $Mode'_3$ using second transmission mode information $Mode'_2$ transmitted from communication terminal apparatus 1050 through transmission path 110 from Equation 5 below and outputs it to signal coding section 102 and signal decoding
25 section 103.

[0135]

[Equation 5]

$$Mode'_3 = \begin{cases} bitrate_1 & (Mode'_1 = bitrate_{low}) \text{ and } (Mode'_2 = bitrate_{high}) \\ bitrate_2 & \left\{ \begin{array}{l} ((Mode'_1 = bitrate_{high}) \text{ and } (Mode'_2 = bitrate_{high})) \\ or \\ ((Mode'_1 = bitrate_{low}) \text{ and } (Mode'_2 = bitrate_{low})) \end{array} \right. \\ bitrate_3 & (Mode'_1 = bitrate_{high}) \text{ and } (Mode'_2 = bitrate_{low}) \end{cases} \dots (5)$$

This is the explanation of the internal configuration of transmission mode determining section 1001 in FIG.10.

[0136] The configuration of transmission mode determining section 1051 in FIG.10 is the same as the configuration of transmission mode determining section 1001 in FIG.10.

[0137] In this way, when there are sounds of running cars or trains or the like on the receiving side, the receiving side recognizes such ambient noise and uses a masking effect of ambient noise and the transmitting side can thereby communicate a speech/audio signal using a minimum transmission bit rate within a range that does not influence human auditory sense and thereby substantially improve the channel efficiency. Furthermore, by detecting not only ambient noise on the receiving side but also information on ambient noise on the transmitting side and using this for coding of a speech/audio signal, it is possible to realize a more efficient communication.

[0138] (Embodiment 3)

Embodiment 3 will explain an example where a

transmission mode information determining method of the present invention is applied to one-way communication typified by music delivery service using portable terminals such as cellular phones.

5 [0139] FIG.12 is a block diagram showing the configuration of a communication apparatus according to Embodiment 3. In FIG.12, communication apparatus 1200 is a communication terminal apparatus on the user side that receives a music delivery service and communication
10 apparatus 1250 is a base station apparatus on the music delivery server side.

[0140] Communication apparatus 1200 is mainly constructed of transmission mode determining section 1201 and signal decoding section 1202. Communication
15 apparatus 1250 is provided with signal coding section 1251.

[0141] Transmission mode determining section 1201 detects ambient noise included in the background of an input signal which is a speech/audio signal, determines
20 a transmission mode for controlling a transmission bit rate at communication apparatus 1250 according to the level of ambient noise and outputs this as transmission mode information to transmission path 110 and signal decoding section 1202.

25 [0142] Signal coding section 1251 performs coding on the input signal based on the transmission mode information transmitted through transmission path 110 and then

integrates it with the transmission mode information and outputs this as coded information to transmission path 110.

[0143] Signal decoding section 1202 decodes coded
5 information transmitted through transmission path 110 and outputs the obtained decoded signal as an output signal. Signal decoding section 1202 compares the transmission mode information included in the coded information output from transmission path 110 with the transmission mode
10 information obtained from transmission mode determining section 1201 with a transmission delay taken into consideration, and can thereby detect transmission errors. To be more specific, when the transmission mode information obtained from transmission mode determining
15 section 1201 with a transmission delay taken into consideration is different from the transmission mode information included in the coded information output from transmission path 110, signal decoding section 1202 decides that a transmission error has occurred in
20 transmission path 110. Furthermore, it is also possible to adopt a technique whereby signal coding section 1251 of communication apparatus 1250 does not integrate the transmission mode information with the coded information, while signal decoding section 1202 decodes the coded
25 information output from transmission path 110 using transmission mode information obtained from transmission mode determining section 1201.

[0144] The internal configurations of transmission mode determining section 1201, signal coding section 1202 and signal decoding section 1251 in FIG.12 are the same as those of transmission mode determining section 101, 5 signal coding section 102 and signal decoding section 103 shown in FIG.2, and therefore detailed explanations of those configurations will be omitted.

[0145] Thus, according to this embodiment, ambient noise in a communication apparatus is detected even in a one-way 10 communication system such as music delivery service and transmission mode information is determined using an auditory masking effect of ambient noise, and therefore base station apparatus can communicate a speech/audio signal using a minimum transmission bit rate within a 15 range that does not influence human auditory sense, and can thereby substantially improve the channel efficiency.

[0146] (Embodiment 4)

Embodiment 4 will explain a case where a transmission mode is determined by decoding coded information 20 transmitted from another party and detecting ambient noise included in the obtained decoded signal.

[0147] FIG.13 is a block diagram showing the configuration of a communication terminal apparatus according to Embodiment 4. In communication terminal 25 apparatuses 1300, 1350 shown in FIG.13, components common to communication terminal apparatuses 100 and 150 shown in FIG.2 are assigned the same reference numerals as those

in FIG.2 and explanations thereof will be omitted.

[0148] When communication terminal apparatus 1300 in FIG.13 is compared to communication terminal apparatus 100 in FIG.2, the operation of transmission mode determining section 1301 is different from that of transmissionmode determining section 101. Furthermore, when communication terminal apparatus 1350 in FIG.13 is compared to communication terminal apparatus 150 in FIG.2, the operation of transmission mode determining section 1351 is different from that of transmission mode determining section 151.

[0149] Transmission mode determining section 1301 detects ambient noise included in a decoded signal, determines a transmission mode for controlling a transmission bit rate when performing coding according to the level of ambient noise and outputs transmission mode information indicating the determined transmission mode to signal coding section 102.

[0150] Next, the internal configuration of transmission mode determining section 1301 in FIG.13 will be explained using FIG.14. Transmission mode determining section 1301 is mainly constructed of masking level calculation section 1401 and transmission mode decision section 1402. As in the case of transmission mode determining section 101 in FIG.2, in addition to a technique of carrying out processing of deciding and outputting the level of ambient noise every time each frame is processed, transmission

mode determining section 1301 in FIG.13 can also perform subsequent processing with pressing of a button by the user of a communication terminal as a trigger or perform subsequent processing at certain time intervals.

5 [0151] As in the case of masking level calculation section 301 in FIG.3, masking level calculation section 1401 calculates the masking level from the decoded signal output from signal decoding section 103 and outputs the calculated masking level to transmission mode decision
10 section 1402.

[0152] As in the case of transmission mode decision section 302 in FIG.3, transmission mode decision section 1402 compares the masking level output from masking level calculation section 1401 with a predetermined threshold,
15 determines a transmission mode for controlling a transmission bit rate based on the comparison result and outputs transmission mode information indicating the determined transmission mode to signal coding section 102.

20 [0153] The internal configuration of transmission mode determining section 1351 in FIG.13 is the same as the configuration of transmission mode determining section 1301, and therefore detailed explanations thereof will be omitted.

25 [0154] Thus, according to this embodiment, by decoding coded information transmitted from the communicating party and detecting ambient noise included in the obtained

decoded signal, it is possible to use the masking effect of ambient noise thereof and perform highly efficient signal coding.

[0155] (Embodiment 5)

5 Embodiment 5 will explain a case where a transmission mode is determined using not only ambient noise on the receiving side included in a decoded signal but also ambient noise on the transmitting side.

[0156] FIG.15 is a block diagram showing the
10 configuration of a communication terminal apparatus according to Embodiment 5. In communication terminal apparatuses 1500 and 1550 shown in FIG.15, components common to those of communication terminal apparatuses 100 and 150 shown in FIG.2 are assigned the same reference
15 numerals as those in FIG.2 and explanations thereof will be omitted.

[0157] When communication terminal apparatus 1500 in FIG.15 is compared to communication terminal apparatus 100 in FIG.2, the operation of transmission mode
20 determining section 1501 differs from that of transmissionmode determining section 101. Furthermore, when communication terminal apparatus 1550 in FIG.15 is compared to communication terminal apparatus 150 in FIG.2, the operation of transmission mode determining section
25 1551 differs from that of transmission mode determining section 151.

[0158] Transmission mode determining section 1501

detects ambient noise included in the background of a speech/audio signal of an input signal, detects ambient noise included in the decoded signal, determines a transmission mode for controlling a transmission bit rate
5 when performing coding according to the level of ambient noise and outputs transmission mode information indicating the determined transmission mode to signal coding section 102.

[0159] Next, the internal configuration of transmission mode determining section 1501 in FIG.15 will be explained
10 using FIG.16. Transmission mode determining section 1501 is mainly constructed of masking level calculation section 1601 and transmission mode decision section 1602. As in the case of transmission mode determining section
15 101 in FIG.2, transmission mode determining section 1501 in FIG.15 can use a technique of performing not only processing of deciding and outputting the level of ambient noise every time each frame is processed but also subsequent processing with pressing of a button by the
20 user of a communication terminal as a trigger or subsequent processing at predetermined intervals.

[0160] Masking level calculation section 1601 calculates a masking level from an input signal and a decoded signal output from signal decoding section 103
25 and outputs the calculated masking level to transmission mode decision section 1602.

[0161] As in the case of transmission mode decision

section 302 in FIG.3, transmission mode decision section 1602 compares the masking level output from masking level calculation section 1601 with a predetermined threshold, determines a transmission mode for controlling a transmission bit rate based on the comparison result and outputs transmission mode information indicating the determined transmission mode to signal coding section 102.

[0162] Here, the processing of masking level calculation section 1601 and transmission mode decision section 1602 will be explained when a method whereby transmission mode determining section 1501 calculates a maximum value and minimum value of the power value of the input signal for a predetermined period, decides the level of ambient noise included in the input signal from the maximum value and minimum value and controls the bit rate according to the level is adopted.

[0163] Masking level calculation section 1601 intervals the input signal into groups of N samples (N : natural number), regards each interval as 1 frame and performs processing in frame units. Hereinafter, the input signal to be coded will be expressed as u'_n ($n=0, \dots, N-1$).

[0164] Furthermore, masking level calculation section 1601 includes buffers $bufu'_i$ ($i=0, \dots, N_i-1$).

[0165] Next, masking level calculation section 1601 will calculate frame power $P_{frameu'}$ of the frame to be processed from Equation 6 below:

[0166]

[Equation 6]

$$P_{frameu'} = \sum_{n=0}^{N-1} |u'_n|^2 \quad \dots (6)$$

Next, masking level calculation section 1601
5 substitutes frame power $P_{frameu'}$ calculated from Equation 6 into buffer $bufu'_{Ni-1}$.

[0167] Next, masking level calculation section 1601
calculates minimum value $P_{frameu'}_{MIN}$ and maximum value
 $P_{frameu'}_{MAX}$ of frame power $P_{frameu'}$ in an i interval
10 (interval length N_i) and outputs $P_{frameu'}_{MIN}$, $P_{frameu'}_{MAX}$
to transmission mode decision section 1602.

[0168] Next, masking level calculation section 1601
updates buffer $bufu'_i$ according to Equation 7 below:

[0169]

15 [Equation 7]

$$bufu'_i = bufu'_{i+1} \quad (i = 0, \dots, N_i - 2) \quad \dots (7)$$

Next, masking level calculation section 1601
intervals the decoded signal output from signal decoding
section 103 into groups of N samples (N : natural number),
20 regards N samples as 1 frame and performs processing in
frame units. Hereinafter, the signal to be coded will
be expressed as decoded signal u''_n ($n=0, \dots, N-1$).

[0170] Furthermore, masking level calculation section
1601 includes buffer $bufu''_i$ ($i=0, \dots, N_i-1$).

25 [0171] Next, masking level calculation section 1601 will
calculate frame power $P_{frameu''}$ to be processed from

Equation 8 below:

[0172]

[Equation 8]

$$P_{frameu}'' = \sum_{n=0}^{N-1} |u''_n|^2 \cdot \cdot \cdot (8)$$

5 Next, masking level calculation section 1601 substitutes frame power P_{frameu}'' calculated from Equation 8 into buffer $bufu''_{N_i-1}$.

[0173] Next, masking level calculation section 1601 calculates minimum value $P_{frameu}''_{MIN}$ and maximum value
10 $P_{frameu}''_{MAX}$ of frame power P_{frameu}'' in an i interval (interval length N_i) and outputs $P_{frameu}''_{MIN}$, $P_{frameu}''_{MAX}$ to transmission mode decision section 1602.

[0174] Next, masking level calculation section 1601 updates buffer $bufu''_i$ according to Equation 9 below:

15 [0175]

[Equation 9]

$$bufu''_i = bufu''_{i+1} \quad (i = 0, \dots, N_i - 2) \quad \dots (9)$$

This is the explanation of the processing by masking level calculation section 1601 in FIG.16.

20 [0176] Next, the processing of transmission mode decision section 1602 will be explained. Transmission mode decision section 1602 determines transmission mode information $Modeu'_1$ from $P_{frameu}''_{MIN}$, $P_{frameu}''_{MAX}$ output from masking level calculation section 1601 according
25 to Equation 10 below:

[0177]

[Equation 10]

$$Modeu'_1 = \begin{cases} \text{bitrate}_{high} & (Thu'_0 \leq Pframeu'_{MAX} / Pframeu'_{MIN}) \\ \text{bitrate}_{low} & (Pframeu'_{MAX} / Pframeu'_{MIN} < Thu'_0) \end{cases} \dots (10)$$

where Thu'_0 is a constant predetermined by an
 5 experiment similar to the aforementioned preliminary
 experiment based on a auditory masking effect of ambient
 noise.

[0178] Next, transmission mode decision section 1602
 determines transmission mode information $Modeu'_2$ from
 10 $Pframeu''_{MIN}$, $Pframeu''_{MAX}$ output from masking level
 calculation section 1601 according to Equation 11 below:

[0179]

[Equation 11]

$$Modeu'_2 = \begin{cases} \text{bitrate}_{high} & (Thu''_0 \leq Pframeu''_{MAX} / Pframeu''_{MIN}) \\ \text{bitrate}_{low} & (Pframeu''_{MAX} / Pframeu''_{MIN} < Thu''_0) \end{cases} \dots (11)$$

15 where Thu''_0 is a constant predetermined by an
 experiment similar to the aforementioned preliminary
 experiment based on the auditory masking effect of ambient
 noise.

[0180] Next, transmission mode decision section 1602
 20 calculates transmission mode information $Modeu'_3$ using
 transmission mode information $Modeu'_1$ and transmission
 mode information $Modeu'_2$ according to Equation 12 below
 and outputs it to signal coding section 102.

[0181]

25 [Equation 12]

$$Modeu'_3 = \begin{cases} bitrate_1 & (Modeu'_1 = bitrate_{low}) \text{ and } (Modeu'_2 = bitrate_{high}) \\ bitrate_2 & \begin{cases} ((Modeu'_1 = bitrate_{high}) \text{ and } (Modeu'_2 = bitrate_{high})) \\ or \\ ((Modeu'_1 = bitrate_{low}) \text{ and } (Modeu'_2 = bitrate_{low})) \end{cases} \\ bitrate_3 & (Modeu'_1 = bitrate_{high}) \text{ and } (Modeu'_2 = bitrate_{low}) \end{cases} \quad \cdot \cdot \quad (12)$$

This is the explanation of the internal configuration of transmission mode determining section 1501 in FIG.15.

5 [0182] The internal configuration of transmission mode determining section 1551 in FIG.15 is the same as that of transmission mode determining section 1501, and therefore explanations thereof will be omitted.

[0183] Thus, according to this embodiment, when there
 10 are sounds of running cars and trains on the receiving side, the transmitting side recognizes ambient noise included in a speech/audio signal transmitted from the receiving side, uses a masking effect of ambient noise and the transmitting side can thereby carry out
 15 communication using a minimum transmission bit rate within a range that does not influence human auditory sense and thereby substantially improve the channel efficiency. Furthermore, by detecting not only ambient noise on the receiving side but also information on ambient
 20 noise on the transmitting side and using it for speech/audio signal coding, it is possible to realize a more efficient communication.

[0184] (Embodiment 6)

Embodiment 6 will explain a case where a relay station

in transmission path 110 adjusts a transmission bit rate transmitted from each communication terminal apparatus in an environment in which communication is carried out according to a scalable coding scheme.

5 [0185] FIG.17 is a block diagram showing the configuration of a communication terminal apparatus and relay station according to Embodiment 6 of the present invention. Furthermore, relay station 1730 exists in midstream of a communication of communication terminal
10 apparatuses 1700 and 1750 in FIG.17. In communication terminal apparatuses 1700, 1750 shown in FIG.17, components common to those of communication terminal apparatuses 100 and 150 shown in FIG.2 are assigned the same reference numerals as those in FIG.2 and explanations
15 thereof will be omitted.

[0186] When communication terminal apparatus 1700 in FIG.17 is compared to communication terminal apparatus 100 in FIG.2, the operations of transmission mode determining section 1701 and signal coding section 1702
20 differ from those of transmission mode determining section 101 and signal coding section 102. Furthermore, when communication terminal apparatus 1750 in FIG.17 is compared to communication terminal apparatus 150 in FIG.2, the operations of transmission mode determining section
25 1751 and signal coding section 1752 differ from those of transmission mode determining section 151 and signal coding section 152.

[0187] Transmission mode determining section 1701 detects ambient noise included in the background of a speech/audio signal in an input signal, determines a transmission mode for controlling a transmission bit rate when performing coding according to the level of ambient noise and outputs transmission mode information indicating the determined transmission mode to transmission path 110 and signal decoding section 103. As in the case of transmission mode determining section 101 in FIG.2, in addition to the technique whereby transmission mode determining section 1701 in FIG.17 performs processing of deciding and outputting the level of ambient noise every time each frame is processed, it is also possible to perform subsequent processing with pressing of a button by the user of the communication terminal as a trigger or perform subsequent processing at predetermined intervals.

[0188] Signal coding section 1702 is fed the input signal and initial transmission mode information, performs coding on the input signal according to the initial transmission mode information and outputs the coded information obtained to transmission path 110. The internal configuration of signal coding section 1702 corresponds to signal coding section 102 shown in FIG.4 with the transmission mode information replaced by the initial transmission mode information.

[0189] Transmission mode determining section 1751

detects ambient noise included in the background of a speech/audio signal in the input signal, determines a transmission mode for controlling a transmission bit rate when performing coding according to the level of ambient noise and outputs transmission mode information indicating the determined transmission mode to transmission path 110 and signal decoding section 153.

[0190] Signal coding section 1752 is fed the input signal and initial transmission mode information, performs coding on the input signal according to initial transmission mode information, integrates an information source code obtained with the initial transmission mode information and outputs this as coded information to transmission path 110.

[0191] Suppose initial transmission mode information mode A in communication terminal apparatuses 1700, 1750 is expressed by Equation 13 below:

[0192]

[Equation 13]

$$ModeA = \begin{cases} bitrate_1 \\ bitrate_2 \\ bitrate_3 \end{cases} \quad \cdot \cdot \cdot (13)$$

The internal configuration of transmission mode determining section 1751 in FIG.17 is the same as that of transmission mode determining section 1701, and therefore explanations thereof will be omitted.

[0193] Next, the internal configuration of relay station

1730 will be explained using FIG.18. In FIG.18, a case where the transmission bit rate of the coded information from communication terminal apparatus 1700 is controlled according to the transmission mode information from communication terminal apparatus 1750 will be explained, but the same applies to a case where the transmission bit rate of the coded information from communication terminal apparatus 1750 is controlled according to the transmission mode information from communication terminal apparatus 1700.

[0194] Relay station 1730 is mainly constructed of interface section 1801, coded information analysis section 1802, transmission mode conversion section 1803, coded information integration section 1804 and interface section 1805.

[0195] Interface section 1801 is fed information transmitted from communication terminal apparatus 1700 through transmission path 110 and transmits information to communication terminal apparatus 1750 through transmission path 110.

[0196] Coded information analysis section 1802 analyzes the information transmitted from communication terminal apparatus 1700, separates it into an information source code and initial transmission mode information mode A coded in their respective layers inside signal coding section 1702 and outputs the information to transmission mode conversion section 1803.

[0197] Transmission mode conversion section 1803 performs transmission bit rate conversion processing on the information source code and initial transmission mode information mode A according to transmission mode information mode B transmitted from communication terminal apparatus 1750. To be more specific, when initial transmission mode information mode A is bitrate 1 and transmission mode information mode B is bitrate 2, transmission mode conversion section 1803 changes initial transmission mode information mode A to bitrate 2 and outputs the base layer information source code, first enhancement layer information source code and initial transmission mode information mode A to coded information integration section 1804. Furthermore, when initial transmission mode information mode A is bitrate 1 and transmission mode information mode B is bitrate 3, transmission mode conversion section 1803 changes initial transmission mode information mode A to bitrate 3 and outputs the base layer information source code and initial transmission mode information mode A to coded information integration section 1804. Furthermore, when transmission mode information mode A is bitrate 2 and transmission mode information mode B is bitrate 3, transmission mode conversion section 1803 changes initial transmission mode information mode A to bitrate 3 and outputs the base layer information source code and initial transmission mode information mode A to coded information

integration section 1804. Furthermore, for combinations of initial transmission mode information mode A and transmission mode information mode B other than those described above, transmission mode conversion section 5 1803 outputs the information source code and initial transmission mode information mode A to coded information integration section 1804 as they are.

[0198] Coded information integration section 1804 is fed the information source code and initial transmission mode 10 information mode A obtained from transmission mode conversion section 1803, integrates them and outputs the integration result as converted coded information to interface section 1805.

[0199] Interface section 1805 is fed information 15 transmitted from communication terminal apparatus 1750 through transmission path 110 and transmits information to communication terminal apparatus 1700 through transmission path 110.

[0200] This is the explanation of the configuration of 20 relay station 1730 in FIG.17.

[0201] Thus, according to this embodiment, when there is ambient noise such as sounds of running cars and trains on the receiving side, the relay station can also control the transmission bit rate instead of the transmitting 25 side. This allows more flexible control of the transmission bit rate and can further improve channel efficiency.

[0202] In this embodiment, the relay station can also determine a transmission mode for controlling a transmission bit rate using not only ambient noise on the receiving side but also ambient noise on the transmitting side.

[0203] FIG.19 is a block diagram showing the configuration of relay station 1730 in this case and the operation of transmission mode conversion section 1901 is different from that of transmission mode conversion section 1803 in FIG.18. Transmission mode conversion section 1901 performs transmission bit rate conversion processing on an information source code and initial transmission mode information mode A according to transmission mode information mode A' and transmission mode information mode B from communication terminal apparatus 1700. To be more specific, when initial transmission mode information mode A is bitrate 1, transmission mode information mode B is bitrate_{high} and transmission mode information mode A' is bitrate_{high}, transmission mode conversion section 1901 changes initial transmission mode information mode A to bitrate 2 and outputs base layer information source code, first enhancement layer information source code and initial transmission mode information mode A to coded information integration section 1804. Furthermore, when initial transmission mode information mode A is bitrate 1, transmission mode information mode B is bitrate_{low} and

transmission mode information mode A' is $\text{bitrate}_{\text{low}}$,
transmissionmode conversion section 1901 changes initial
transmission mode information mode A to bitrate 2 and
outputs the base layer information source code, first
5 enhancement layer information source code and initial
transmission mode information mode A to coded information
integration section 1804. Furthermore, when initial
transmission mode information mode A is bitrate 1,
transmission mode information mode B is $\text{bitrate}_{\text{low}}$ and
10 transmission mode information mode A' is $\text{bitrate}_{\text{high}}$,
transmissionmode conversion section 1901 changes initial
transmission mode information mode A to bitrate 3 and
outputs base layer information source code and initial
transmission mode information mode A to coded information
15 integration section 1804. Furthermore, when initial
transmission mode information mode A is bitrate 2,
transmission mode information mode B is $\text{bitrate}_{\text{low}}$ and
transmission mode information mode A' is $\text{bitrate}_{\text{high}}$,
transmissionmode conversion section 1901 changes initial
20 transmission mode information mode A to bitrate 3 and
outputs the base layer information source code and
transmission mode information mode A to coded information
integration section 1804. Furthermore, for combinations
of initial transmission mode information mode A,
25 transmission mode information mode B and transmission
mode information mode A' other than those described above,
transmission mode conversion section 1901 outputs the

information source code and transmission mode information mode A to coded information integration section 1804 as they are.

[0204] Thus, according to this embodiment, when there
5 is ambient noise such as sounds of running cars and trains on the receiving side and transmitting side, the relay station can also control the transmission bit rate instead of the transmitting side. This allows more flexible control of the transmission bit rate and can further
10 improve channel efficiency.

[0205] When a certain relay station exists in transmission path 110 in an environment in which a communication of a speech/audio signal under a one-way communication scheme is being carried out according to
15 a scalable coding scheme, combining this embodiment with above described Embodiment 3 will also allow the relay station to use transmission mode information transmitted from the communication terminal, reduce the amount of information of the coded information transmitted from
20 the base station and retransmit it to transmission path 110.

[0206] The present application is based on Japanese Patent Application No.2004-048569 filed on February 24, 2004, entire content of which is expressly incorporated
25 by reference herein.

Industrial Applicability

[0207] The present invention is suitable for use in a communication terminal apparatus of a packet communication system or mobile communication system.